

Hyperion Project Begins Collaboration to Develop an Advanced Technology Cluster Test Bed

SIGNIFICANCE

One of the challenges associated with developing, managing, and utilizing high-performance computing (HPC) systems is a lengthy and costly testing phase that must be completed to ensure the new hardware, middleware, and application software works well at scale. LLNL has teamed with 10 industry partners—Cisco Systems, DataDirect Networks, Dell, Intel, LSI, Mellanox, QLogic, Red Hat, Sun Microsystems, and Supermicro—to develop an innovative solution to this problem. Hyperion is a large-scale Linux cluster test environment, which will speed the development and reduce the cost of powerful HPC clusters by making a petascale cluster available solely for development and testing infrastructural components. It will also be used to evaluate system hardware and software critical to maintaining the aging U.S. nuclear weapons stockpile. Hyperion, too large for any one organization to deploy, represents an innovative approach for industry-leading collaborations and exemplifies LLNL's leadership in HPC.

PROGRESS IN 2008

Michael Dell's announcement of the Hyperion Project during his 2008 Supercomputing Conference (SC08) keynote address was the culmination of more than a year and a half of work between LLNL and industry collaborators to field an advanced technology cluster as a testing and scaling resource. "Storage, connectivity, management software: these are all challenges that we are going to be dealing with as we implement [hyperscale] systems," Dell said. "Hyperion is a test bed big enough to really test scale, and it will share those breakthroughs with the entire open-source community."

Hyperion is a dedicated development and test environment for advanced Linux cluster technologies, an evaluation

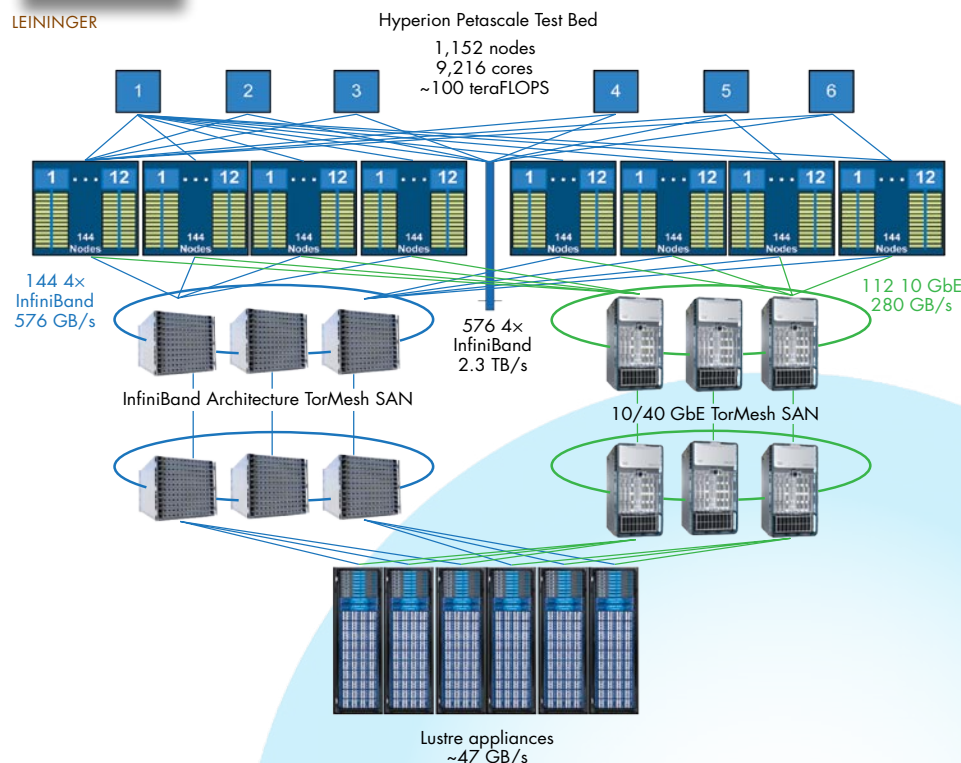
test bed for new hardware and software technologies, and an input and output (I/O) test bed to prepare for LLNL's future ASC Sequoia platform. The first phase of Hyperion was installed at LLNL in September 2008 and consists of 576 nodes comprised of two-socket quad-core Intel Harpertown processors, eight gigabytes (GB) of memory per node, and a double data rate (DDR) InfiniBand interconnect. The first-phase cluster is testing the Lustre parallel file system and the Red Hat-based ASC Tripod Operating System Software (TOSS). It is also testing Message Passing Interface (MPI), OpenMP, and the OpenFabrics high-performance networking software for InfiniBand and low-latency Ethernet. DataDirect Networks, LSI, and Sun Microsystems each provided high-



LEININGER

CONTACT INFORMATION

MATT LEININGER,
MATT@LLNL.GOV



The Hyperion cluster architecture includes eight scalable units (1,152 nodes), the InfiniBand and Ethernet TorMesh SAN, and high-performance storage. Hyperion connects to one SAN via 144 4x-InfiniBand network links (blue lines) and to the other SAN by 112 10-GbE links (green lines). There are also 576 4x-InfiniBand links, totaling 2.3 terabytes (TB) of bandwidth, that are not currently in use.

performance storage for Hyperion, which is connected through two separate storage area networks (SANs), one based on DDR InfiniBand and another based on low-latency 10-gigabit Ethernet (GbE).

The second phase of Hyperion, scheduled for March 2009, will double the system size to 1,152 nodes, with the

additional nodes using the new Intel Nehalem processors (two-socket quad-core), 12 GB of memory per node, and a DDR InfiniBand interconnect. All Hyperion nodes have Peripheral Component Interconnect Express 2.0, which enables future upgrades to quad data rate (QDR) InfiniBand.